

AD-A125 134

SPEECH RECOGNITION: ACOUSTIC PHONETIC AND LEXICAL
KNOWLEDGE REPRESENTATION(U) MASSACHUSETTS INST OF TECH
CAMBRIDGE RESEARCH LAB OF ELECTRONICS V W ZUE

1/1

UNCLASSIFIED

22 FEB 83 N00014-82-K-0727

.F/G 5/8

NL

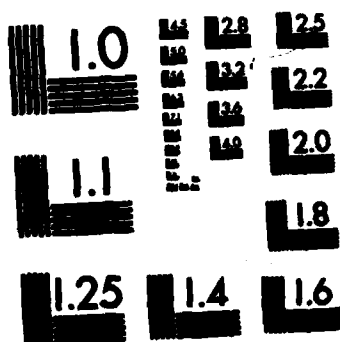


END

FILMED

+

010



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

AD A125134

STATUS REPORT

SPEECH RECOGNITION: ACOUSTIC PHONETIC
& LEXICAL KNOWLEDGE REPRESENTATION

OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY

Contract Number N00014-82-K-0727

covering the periods

July 1, 1982 - September 30, 1982
October 1, 1982 - December 31, 1982

Submitted by
Victor W. Zue

February 22, 1983

DTIC
ELECTE
S MAR 2 1983 D
D

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

DTIC FILE COPY

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

88 02 028 170

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

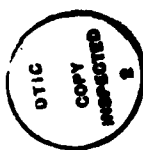
REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A125134	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SPEECH RECOGNITION: ACOUSTIC PHONETIC AND LEXICAL KNOWLEDGE REPRESENTATION		5. TYPE OF REPORT & PERIOD COVERED Status Report 07/01/82 - 09/30/82 10/01/82 - 12/31/82
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Victor W. Zue		8. CONTRACT OR GRANT NUMBER(s) N00014-82-K-0727
9. PERFORMING ORGANIZATION NAME AND ADDRESS Research Laboratory of Electronics Massachusetts Institute of Technology Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-049-542
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research - Dept. of the Navy - 800 North Quincy Street Arlington, Virginia 22217		12. REPORT DATE February 22, 1983
		13. NUMBER OF PAGES 5
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) A large vocabulary and a continuous speech recognition system, which are acoustic-phonetically based phontactic constraints acoustic cues to word boundaries. <i>→ the purpose of this program is to</i>		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Develop a speech data base facility under which the acoustic characteristics of speech sounds in various contexts can be studied conveniently; investigate the phonological properties of a large lexicon of, say 10,000 words, and determine to what extent the phonotac- tic constraints can be utilized in speech recognition; study the acoustic cues that are used to mark work boundaries; develop a test bed in the form of a large- vocabulary, IWR system to study the interactions of acoustic, phonetic and lexi- cal knowledge; and develop a limited continuous speech recognition system with the goal of recognizing any English word from its spelling in order to assess the in-		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. teractions of higher-level knowledge sources.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



PROGRESS REPORT

Contract Number N00014-83-K-0727

Periods Covered: 07/01/82 - 09/30/82
10/01/82 - 12/31/82

Implementation of the essential parts of the SPIRE speech research facility has been completed. Speech signals can now be digitized, analyzed and displayed conveniently. We have now digitized approximately thirty minutes of speech spoken by seven different speakers. Phonetic transcriptions for some of these words and sentences have also been entered. In addition, SPIREX, a data base facility to examine the statistical properties of speech sounds, has also been implemented on the Lisp machine.

A study of the structural constraints on the sound patterns in American English has also been completed. With the help of an on-line dictionary, we were able to determine the various ways with which words can be accessed without detailed phonetic information. The 20,000 word Merriam-Webster Pocket Dictionary has been carefully scrutinized in terms of the phonemic spellings and stress markers.

In order to design the corpus to be used to study the acoustic cues to word boundaries, the Harvard phonetically balanced sentences were carefully examined. Our initial conclusion is that the distribution of consonants with words and across word boundaries are sufficiently rich that the Harvard sentences can be used as the initial corpus. Data for the pilot study were re-examined using the SPIREX facility for confirmation.

A preliminary version of the "front-end" portion of the large vocabulary isolates word recognition system has been implemented. The system can now obtain, from the digitized waveform, a set of acoustic parameters. From these acoustic parameters, the speech signal is segmented into acoustic regions. A set of re-write rules is then applied to transfer the acoustic segments into a string (or lattice) of broad phonetic labels. This string is used to access a 20,000 word lexicon, and the result is a set of words satisfying the broad phonetic description. So far we have tested approximately 100 words spoken by two male speakers, and the results are promising. Very often, this procedure will result in significant reduction (down to <1%) of the lexical candidates.

A preliminary version of the continuous digit recognition system has been implemented. The system is intended to utilize acoustic phonetic knowledge, and the knowledge is implemented in a modular fashion so that it can easily be extended to other tasks. The current implementation will produce a broad phonetic description of the signal, which can then be used to propose the underlying digit strings. Due to the strong constraints on the task, we expect this initial broad classification will very often produce a unique digit string.

Page 2
PROGRESS REPORT
Contract Number N00014-83-K-0727

A variety of system software has been written on the Lisp machine. This includes device drivers for the Digital Sound A/D-D/A converter and the Versatek Printer Plotter and a critical-band digital fitter bank for spectrum analysis.

DISTRIBUTION LIST

	<u>DODAAD</u>	<u>Code</u>
Director Advanced Research Project Agency 1400 Wilson Boulevard Arlington, Virginia 22209 Attn: Program Manager	HX1241	(2)
Associate Director for Mathematical and Physical Sciences Research Program Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217	N00014	(3)
Administrative Contracting Officer E19-628 Massachusetts Institute of Technology Cambridge, Massachusetts 02139		(1)
Director Naval Research Laboratory Washington, D.C. 20375 Attn: Code 2627	N00173	(6)
Defense Technical Information Center Bldg. 5, Cameron Station Alexandria, Virginia 22314	S47031	(12)
TACTEC Battelle Memorial Institute 505 King Avenue Columbus, Ohio 43201	79986	(1)